基于科研产出关联的大科学装置集群规划布局分析方法研究

冯凌子[1,2]　张汝昊[1,2]　冯凯悦[3]　袁军鹏[1,2]
[1]（中国科学院文献情报中心　北京　100190）
[2]（中国科学院大学经济与管理学院图书情报与档案管理系　北京　100049）
[3]（中国科学院条件保障与财务局　北京　100086）

摘要：

[目的] 大科学装置对促进重大科技成果产出具有重要作用，其集群化发展协同创新效应不断引领区域产业创新发展。为了使大装置更好地发挥协同效应，需要更好地进行大装置规划布局。

[方法] 本研究从科研成果入手，提出一种通用的分析框架，可用于辅助分析大科学装置间学科层面的关联程度，作为以需求驱动规划的补充。以广东松山湖科学城为例，收集已有大装置（散裂中子源与同步辐射光源）与规划布局的大装置 Y 在国际上同类型装置的论文数据，分析其学科层面的内在关联，用以辅助判断规划布局的装置与已有装置的匹配关联程度。

[结果] 已有装置（散裂中子源与同步辐射光源）关联度高，存在可能的互补关系，大装置 Y 与已有装置关联度则较低，如要布局建议还需要进一步专家论证。

[局限]本研究主要从装置产出成果来量化装置间联系，只可为决策者提供一种客观的数据视角和辅助参考，在大科学装置集群的规划方面，更重要的是国家整体的规划和区域对大科学装置的需求，本方法需要融合大科学装置的内在原理、运作流程以及专家知识，才可在规划布局中发挥真正的作用。

[结论] 该方法一定程度上能为大装置规划布局提供前瞻性的定量化参考，从而为大装置群的功能协调与未来发展提供预见性、全局性的优化方案。

关键词：大科学装置，大科学装置集群，大科学装置布局，国家重大科技基础设施，科研产出

分类号：G3

Research on the analysis method of large-scale scientific facilities cluster layout planning based on the correlation of scientific research output

Feng Lingzi[1,2]　Zhang Ruhao[1,2]　Feng Kaiyue[3]　Yuan Junpeng[1,2*]
[1] (National Science Library, Chinese Academy of Sciences, Beijing 100190, China)
[2] (Department of Library, Information and Archives Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100049, China)
[3] (Condition Guarantee and Finance Bureau, Chinese Academy of Sciences, Beijing 100086, China)

[Abstract]

[Objective] Large-scale scientific facilities play an important role in promoting the output of major scientific and technological achievements, and the collaborative innovation effect of its cluster development continues to lead the innovation and development of regional industries. In order to make the large-scale scientific facilities better play the synergistic effect, it is necessary to better plan the layout.

[Methods] This study proposes a general analytical framework based on scientific research results. It can be used to assist in the analysis of the degree of interdisciplinary correlation between large-scale scientific facilities, as a supplement to demand-driven planning. Taking the Songshan Lake Science City in Guangdong as an example, collecting papers on existing large-scale facilities and large facilities of the same type in the world. Analyze the internal correlation of the data at the subject level to assist in judging the degree of matching and correlation between the planned layout of the facility and the existing facility.
[Results] Existing devices have a high degree of correlation, and there is a possible complementary relationship, while the relationship between the large device Y and the existing device is relatively low. If the layout of this facility is to be proposed, further expert argumentation is required.
[Limitations] This research mainly quantifies the connection between large-scale scientific facilities from the output results of the device, and can only provide an objective data perspective and auxiliary reference for decision makers. In the planning of large large-scale scientific facilities clusters, the more important is the overall national planning and the regional needs of large-scale scientific facilities. This method needs to integrate the internal principles, operation procedures and expert knowledge of large-scale scientific facility before it can play a real role in planning and layout.
[Conclusions] The results show that this method can provide a forward-looking quantitative reference for the planning and layout of large-scale facilities to a certain extent, so as to provide a foreseeable and overall optimization plan for the coordination and future development of large-scale facilities.
Keywords: large-scale scientific facility, large-scale scientific facility cluster, large-scale scientific installation layout, Major National Science and Technology Infrastructure, scientific research output

1 引言

大科学装置是为探索未知世界、发现自然规律、实现技术变革提供极限研究手段的大型复杂科学研究系统，是突破科学前沿、解决经济社会发展和国家安全重大科技问题的大规模基础设施[1]。目前，国内外已有越来越多的重大科学发现依托于大科学装置产生[3;4]，同时，大科学装置在整合科技力量与资源、产生科技人才集聚效应、提升国家和地区创新能力和国际竞争力等方面更展现出了重要而独特的作用[2]。自 2009 年以来，我国科技部、中科院提出了国家重大科技基础设施建设中长期规划（2012-2030 年）和 2050 年大科学装置发展路线图等系列规划，指出要面向我国国家战略需求和科技前沿发展需要，系统部署并建成若干世界一流的大科学装置。

近年来，大科学装置个体的领域局限性与愈演愈烈的学科交叉态势间的矛盾推动了国内外大科学装置的集群化发展，由于存在技术互补等特点，集群的大科学装置相互之间的协同创新效应亦日益明显，不断引领区域产业创新发展[2]，例如美国橡树岭国家实验室（Oak Ridge National Laboratory，ORNL）汇聚了高通量同位素反应堆（HFIR）、电子直线加速器（ORELA）、散裂中子源（SNS）、“顶点”超级计算机（Summit）等多个大科学装置，已发展成为大型综合性研究基地，对美国中子科学、先进材料、纳米技术等众多领域的发展做出了巨大贡献；英国哈威尔科学和创新园（Harwell Science and Innovation Campus，HSIC）汇聚了“钻石”同步辐射光源（Diamond）、脉冲散裂中子源（ISIS）、中央激光设施（CLF）、新光源（NLS）等大科学基础设施，在物理、化学、材料科学、地质学、生物医

学、环境等领域云集了大批世界顶尖科学家，成为了集科学、创新和企业为一体的世界领先高科技园区。日本高能加速器研究机构（High Energy Accelerator Research Organization，KEK）汇集了脉冲散裂中子装置（KENS）、光子工厂先进环（PF-AR）、共建质子同步加速器（J-PARC）等大装置，对日本科技竞争力的提升起到了重要作用。我国通过中央与地方共建等方式充分调动地方积极性，在北京、合肥、上海等地相继布局建设了一批大科学装置，已形成大科学装置集群雏形。

大科学装置集群的意义是重大的，其针对产业技术发展中存在的关键和共性问题，能够通过装置间的设施和知识信息共享形成知识协同，实现学科交叉和优势互补，从而在更复杂、微观的层面冲破技术壁垒、实现原始知识创新[2]。此外，大科学装置集群所带来的人才、资源、合作、产业等的“集聚效应”、“标杆效应”[5]，将更有利于带动大规模、大范围的区域经济发展和产业升级。大科学装置集群的潜在效益对于处在建设科技强国关键时期的我国来说更是必要的，“十四五”规划中就指出，要从战略导向性、应用支撑性、前瞻引领性及民生改善性四个作用维度对国家重大科技基础设施进行区域系统布局，从而支撑北京、上海、粤港澳大湾区形成国际科技创新中心，北京怀柔、上海张江、大湾区、安徽合肥建成综合性国家科学中心[6]。

大科学装置耗资巨大，为实现大科学装置集群的实用价值最大化，在符合国家整体规划和区域对大科学装置需求的前提下更需要合理地布局。大科学装置群布局研究在近年正逐步受到国内学界的关注。有学者认为，大科学装置是一种“高密度科技创新资源聚合体”，是粤港澳大湾区应对全球性创新资源竞争的有力支撑，但同时，广东省乃至全国范围内的大科学装置布局也存在过于分散等问题，需要力争集聚更多大科学装置并更合理地布局，以提升创新资源集聚效力[7]。部分学者以依托大科学装置的国际创新科技园区、湾区为调查对象，梳理大科学装置建设现状和发展趋势，从而借鉴大科学装置群的布局经验。如张玲玲等探讨了散裂中子源大科学装置于科学园区的空间布局与功能定位[8]；陈套和冯锋强调了大科学装置群的内涵、优势与集群效应，并呼吁大科学装置集群发展，以激发创新和学科交叉潜力[5]；陈岸明和魏东原则通过比较，总结了粤港澳大湾区内重大科技基础设施建设在统筹管理、人才供给、集群规划等方面的不足[9]。程晓舫等人选取长三角作为典型地区，利用经济建模和演绎法分析大科学装置优化布局的思路及社会效益，研究发现合理应用和功能规划并举、加强大装置集群的共建共管共享、促进区域科技和经济联动发展，将有利于真正发挥大科学装置的作用[10]。也有学者认为学科的布局结构也影响大科学装置科学效益的发挥[11]，例如张玲玲等以上海光源为例分析了大科学装置学科主题与合作网络[11]。相较而言，国外学者对于大科学装置及集群的前期布局研究关注较少，主要关注大科学装置的各方面效益，例如经济社会效益[12; 13]，产业创新效率[14]，产出研究成果绩效（如论文数量[15; 16]、影响因子[17]等），以及人才培养和合作[18]等。

总体来看，当前研究虽然对大科学装置群的规划、布局和可能带来的效益给予了重点关注，但大部分停留于理论探讨，而尚未涉及切实的布局策略；虽然肯定了大科学装置群布局可能激发的学科交叉与科技创新潜力，但却鲜有研究从大科学装置群的学科契合度层面具体展开布局的合理性分析。而目前看来，已有的大科学装置集群在研究上具有学科集聚效应，在学科交叉极大促进创新的态势下，合理地对学科进行规划能够进一步增加学科协作效应、促进创新。因此大科学装置规划建设过程中系统性学科布局是一个亟需解决的现实问题[11]，目前大科学装

置群的实际布局工作，主要以宏观层面的国家需求为导向[5]，在前期项目论证中一定程度上存在“只见树木，不见森林”“只见学科、不见关联”的情况，对载体间、载体和项目间关系的论证不足，前沿引领和综合交叉性体现不够，因此亟待开展相关研究，强化设施间、设施与平台的学科关联度研究，在大学科规划下明确交叉领域布局，形成先行启动区的学科定位特色。

本文的重要产出是为需求驱动的大装置规划提供定量层面的学科布局决策参考建议，内在联系更加紧密、更合理的大科学装置布局能够提高产生的效益，对国民经济产生更显著、更及时的正向影响。为了对大科学装置集群更科学布局在学科层面提供决策参考信息，本研究从科研成果入手，提出一种通用的分析框架，可用于辅助分析大科学装置间学科层面的关联程度，有助于为大装置规划布局提供前瞻性的定量化参考，从而从学科布局上为大装置群的功能协调与未来发展提供优化方案，作为以需求驱动规划的补充。

值得说明的是，本研究方法对大科学装置群的规划有一定参考意义，但大科学装置集群规划更重要的是国家整体的规划和区域对大科学装置的需求。

2 大科学装置集群规划布局分析框架

由于学科布局结构对大科学装置效益产生巨大影响，为了使大科学装置集群能够通过装置之间的协同更大化地发挥作用，那么装置之间应当在学科层面以及研究层面具有足够的内在关联，所以可以通过获取大装置的论文产出数据来表征装置在学科与研究层面的关联情况。

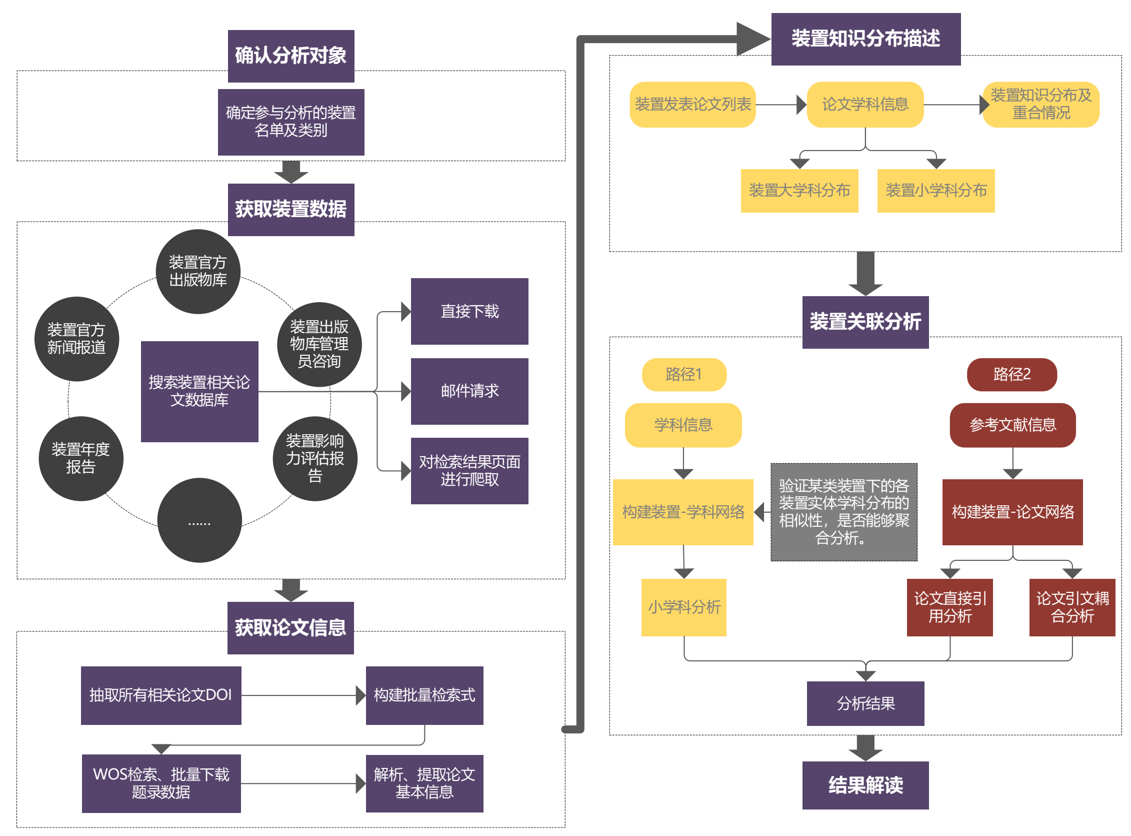


图 1 技术路线

方法的技术路线如图 1 所示，具体步骤如下：

（1）确认分析对象

确定所要分析的大装置对象，如果需要分析已布局的大装置集群间装置的关联，可以通过大装置已有的论文产出来分析表征，如果需要分析正在规划布局的大装置与已有装置之间的关联程度，可以通过与所要分析的大装置对象同类型的已布局大装置产出情况来替代分析关联。

（2）获取装置数据

从各大装置的官方出版物库、官方新闻报道、装置年度报告、装置影响力评估报告，及联系装置出版物库管理员等方式，结合直接下载、邮件请求与爬虫爬取等方法获取依托各个大装置发表论文的数据。

（3）获取论文信息

通过 DOI 等论文标识符在论文题录数据库（如 Web of Science）获取论文题录信息，解析抽取学科、引用关系等信息。

（4）装置知识分布描述

通过装置发表论文学科情况，得到装置的知识分布：

论文通过依托大装置开展研究且发表，同时论文有归属学科，通过这一关系可以得到装置与学科的对应关系，进而得到装置在各个学科的知识分布。

再构建“装置-学科”对应关系，得到被分析装置所涉及的重要学科及累积论文比例，如图 2 所示。装置到学科连线颜色越深，表示该装置这一学科论文占比越高；右侧学科颜色越深，表示该学科是几种装置主要知识来源。

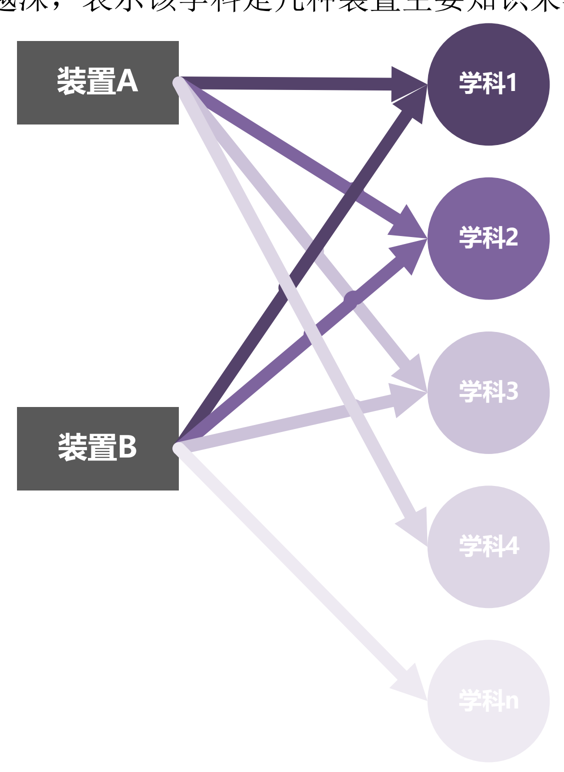


图 2 “装置-学科”知识占比示意图

本研究在学科分析中将同时使用大学科与子学科两种学科粒度，大学科采用 ESI 的学科分类[1]，子学科采用 WoS 学科分类[2]。

[1] ESI 是基于 SCI（科学引文索引）和 SSCI（社会科学引文索引）文献记录而建立的计量分析数据库，该数据库设置了 22 个学科分类。

[2] Web of Science 是科睿唯安旗下的全球学术数据库平台，包含 SCI、SSCI、A&HCI 等数据库，学科分类涵

（5）装置关联分析

本研究通过两种路径共同实现“装置与装置”之间的关联分析：

路径 1：构建“装置-学科”的二分网络，通过学科共现情况和学科分布的相似程度，得到装置两两之间的关联。

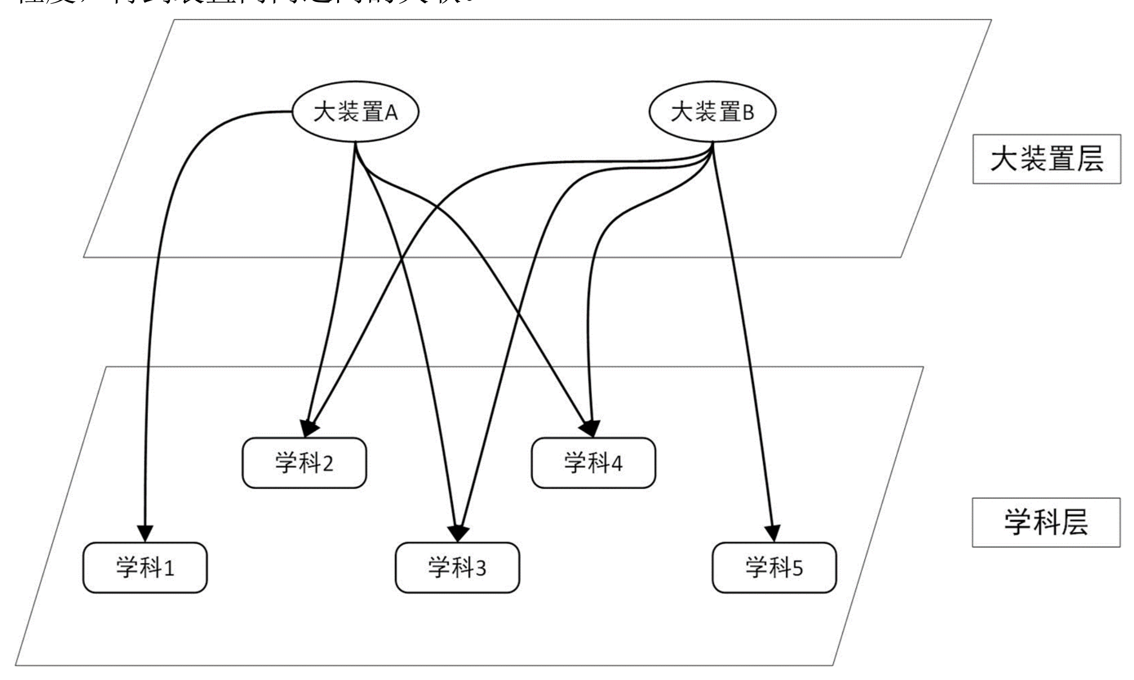


图 3 “装置-学科”网络示意图

如图 3 所示，构建装置层与学科层两层网络，论文通过依托大装置开展研究且发表，同时论文有归属学科，通过这一关系可以得到装置与学科的对应关系，进而得到装置与装置间的学科共现联系。但普通的学科共现联系强度算法可能受到不同大装置论文体量的影响，为了更准确地衡量各大装置间的学科分布的相似程度，可利用大装置学科分布向量的余弦相似度获得，定义为：

$$\mathrm{S}\,(\mathrm{A},\mathrm{B}) = \theta(v_A, v_B) = \frac{\sum_{\mathrm{k}\in\Phi} v_A^k \cdot v_B^k}{\sqrt{\sum_{\mathrm{k}\in\Phi} (v_A^k)^2}\sqrt{\sum_{\mathrm{k}\in\Phi} (v_B^k)^2}} \tag{1}$$

其中， $\Phi = [\varphi_1, \varphi_2, \dots, \varphi_m]$ 表示现有的所有子学科分类，v_A、v_B分别表示大装置 A 和 B 的学科分布向量，例如：$v_A = [v_A^{\varphi_1}, v_A^{\varphi_2}, \dots, v_A^{\varphi_m}]$，$v_A^{\phi_1}$则表示大装置 A 在学科$\varphi_1$上的论文分布数量占 A 所有论文数量的比重。

路径 2：构建“装置-论文”的二分网络，通过论文的关联间接得到装置间的关联

盖 252 个学科。

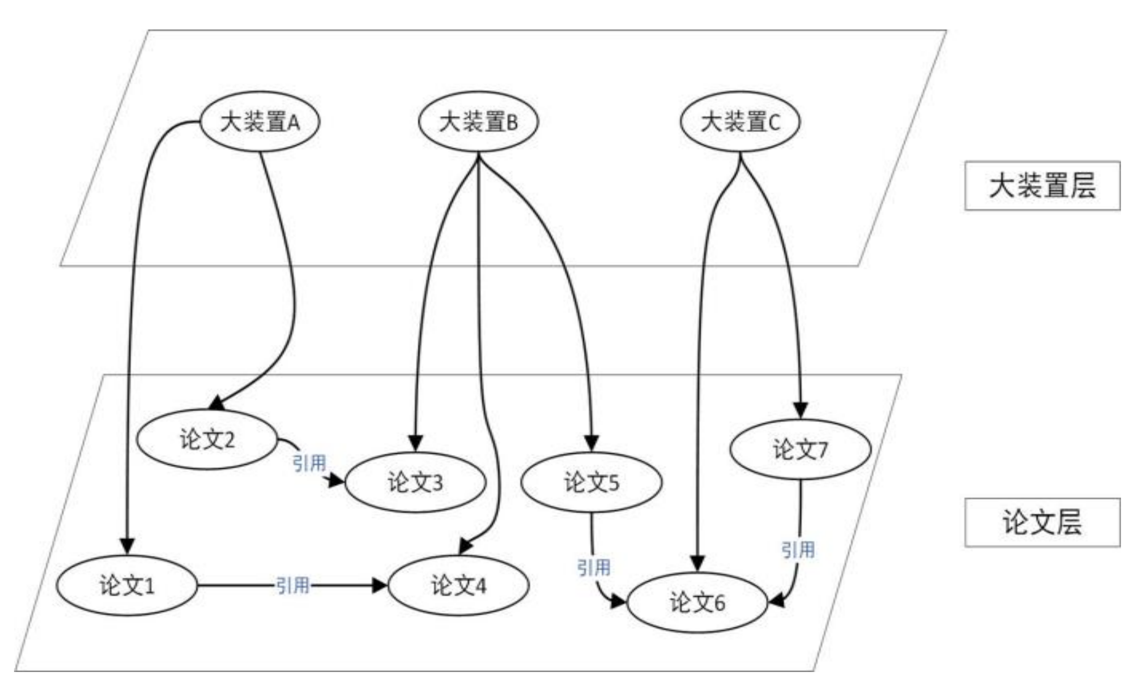


图 4 “装置-论文”网络示意图

如图 4 所示，构建装置层与论文层两层网络，依托装置所发表的论文之间可能存在引用关系，那么装置之间的关联，可以通过论文之间的关联来间接表征。例如，如果存在论文 1 引用了论文 4，论文 2 引用了论文 3，论文 1 与论文 2 通过依托大装置 A 发表，论文 3 与论文 4 通过依托大装置 B 发表，那么在大装置层即构成一条原始权重为 2 的由大装置 A 指向大装置 B 的连边，以此类推。

设P_A、P_B为依托大装置 X 和 B 发表的论文集合，R_A、R_B为P_A、P_B的参考文献集合，那么装置 A 与 B 的论文直接引用关系强度$DCS(A,B)$和论文耦合关系强度$BCS(A,B)$可以分别定义为以下两式。

$$DCS(A,B) = AVG\left(\frac{|R_A \cap P_B|}{|R_A|}, \frac{|R_B \cap P_A|}{|R_B|}\right) \tag{2}$$

$$BCS(A,B) = \frac{|R_A \cap R_B|}{|R_A \cup R_B|} \tag{3}$$

（5）结果解读

根据数据结果，解读分析大装置之间的关联情况。

3 案例与数据获取

本研究以广东省松山湖科学城为例，分析其已有大装置与规划布局的大装置之间的关联。松山湖科学城目前已布局中国散裂中子源大装置，南方先进光源研究测试平台和预研项目也正在有序推进中，松山湖科学城已初具重大科技基础设施集群趋势，对高端创新资源的集聚效应逐步凸显，在此基础上规划新增某大装置 X，为了更好地使大装置发挥协同效应，新规划布局的大装置应在学科层面具有较强的相关性，从而保证装置之间的协同作用。

（1）确定分析的大装置对象

由于松山湖科学城已有大装置产出论文数量目前暂时有限，研究选择了国内

外同类型装置作为替代，获取装置已发表论文数据来进行分析。目前选定作为替代分析的装置见表 1。

散裂中子源装置的替代分析装置选择了国际上三个已建成的散裂中子源，包括美国散裂中子源（SNS）、英国散裂中子源（ISIS），以及日本散裂中子源（J-PARC）。

南方光源的替代分析装置选择上，首先列出了与南方光源同能量等级的已建成光源列表作为备选，其中上海光源是我国成功的第三代同步辐射光源，既是国内光源又是同能量级光源，但上海光源成果的获取渠道和最终获取数据量实在有限，最终仅获取到 127 条已报道成果，直接用来分析与其他设施相比偏差较大，综合考虑了各装置的数据的完备性和可获取性，最终分析装置选择了英国 Diamond 光源和西班牙 ALBA 光源。

规划布局的某大装置 X 的替代分析装置则选择了国际上同类型的大装置 Y。

表 1 松山湖科学城待分析大装置及替代分析同类型大装置

松山湖科学城大装置	替代分析同类型大装置
散裂中子源	美国散裂中子源（SNS）
	英国散裂中子源（ISIS）
	日本散裂中子源（J-PARC）
南方光源	英国Diamond光源
	西班牙ALBA光源
规划布局的某大装置X	国际上同类型大装置Y

（2）获取数据

确定待分析的大装置列表后，研究通过从各大装置的官方出版物库、官方新闻报道、装置年度报告、装置影响力评估报告，及联系装置出版物库管理员等方式，结合直接下载、邮件请求与爬虫爬取等方法获取依托各个大装置发表论文的数据。

利用获取到的大装置论文的论文标识符（DOI），在 Web of Science 数据库中进行匹配，批量获取论文的题录数据。研究数据均获取于 2020 年 8 月。

4 结果与讨论

4.1 装置论文发表情况

依托散裂中子源发表论文情况如图 5 所示，截至 2020 年 8 月，总计发表论文 14384 篇，其中英国散裂中子源 11381 篇，美国散裂中子源 2100 篇，日本散裂中子源 903 篇。论文数量自 1978 起开始波动上升，2004-2009 年呈下降趋势，自 2009 年起出现了明显的上升，并在 2018 年达到峰值 907 篇。

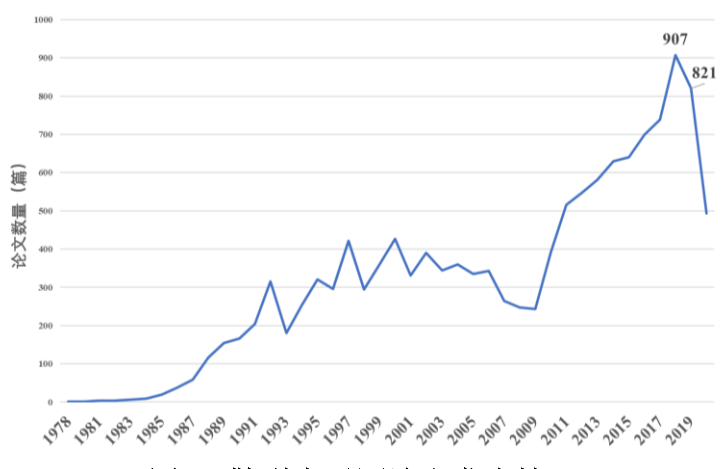


图 5 散裂中子源论文发表情况

依托同步辐射光源发表论文情况如图 6 所示，截至 2020 年 8 月，总计发表论文 10421 篇，其中英国 Diamond 光源 9141 篇，西班牙 ALBA 光源 1280 篇。论文数量自 2001 起开始近乎呈指数上升，在 2017 年左右增速开始放缓，2019 年达到峰值 1398 篇。

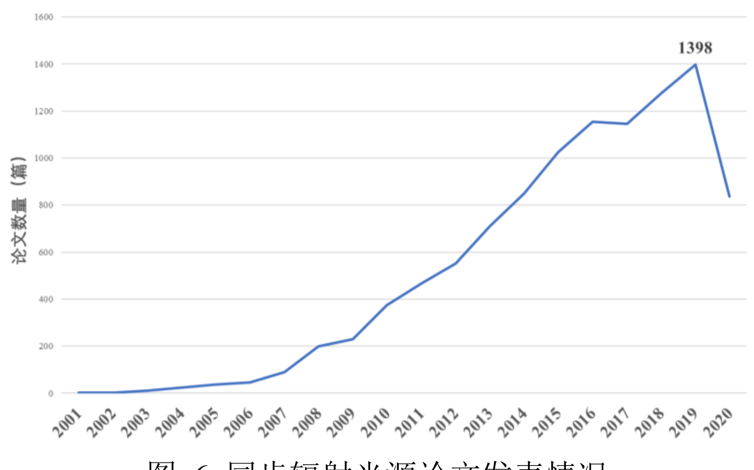


图 6 同步辐射光源论文发表情况

与规划布局的某大装置 X 同类型的大装置发表论文情况如图 7 所示，由于落成时间相对较晚，截至 2020 年 8 月，该装置总计发表论文 247 篇。论文数量自 2013 年起持续上升，在 2018 年达到峰值 66 篇后有所下降。

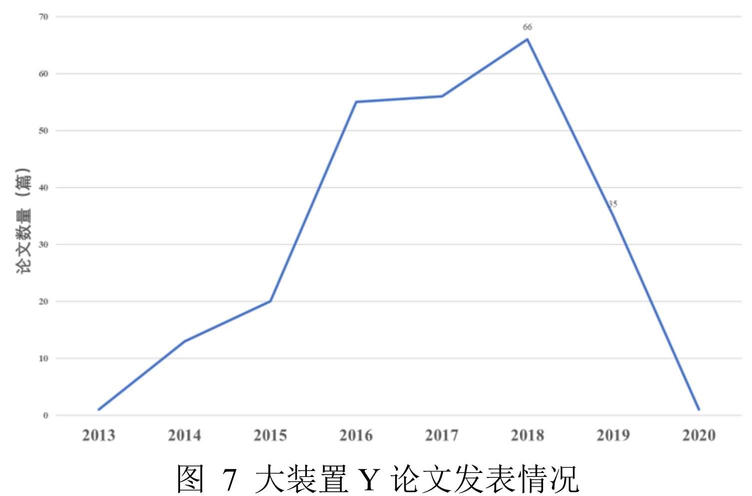


图 7 大装置 Y 论文发表情况

4.2 装置学科分布

为了了解松山湖初建成与在建的三类科学大装置的学科属性，首先对这三类大装置的国际对标装置（包括美国散裂中子源（SNS）、英国散裂中子源（ISIS）、日本散裂中子源（J-PARC）、西班牙光源（ALBA）、英国“钻石”同步辐射光源（DIAMOND）、大装置 Y 进行学科分析。

在大学科（ESI 学科分类）层面的结果如表 2 所示，物理学、化学是三类科学装置的相关论文均涉及最多的两大学科类，占比分别达到了 66.7%、41.8%和78.9%，可见大科学装置对于这类基础学科研究有着不可替代的作用。其次是材料科学和生物与生化，在三类大科学装置的论文中也占有较大比重。此外，三类大科学装置都包含部分交叉科学类研究，这说明三类大科学装置除对多种主要学科具有支撑作用外，也具备支持交叉科学研究的潜力。

表 2 三类科学装置大学科分布 Top 5

	散裂中子源	**同步辐射光源**	**大装置Y**
1	物理学（5324，37.0%）	化学（3070，29.5%）	物理学（168，68%）
2	化学（4275，29.7%）	物理学（1286，12.3%）	化学（27，10.9%）
3	材料科学（1681，11.7%）	交叉科学（1278，12.3%）	交叉科学（15，6.1%）
4	交叉科学（415，2.9%）	生物与生化（1267，12.2%）	材料科学（14，5.7%）
5	生物与生化（207，1.4%）	材料科学（1073，10.3%）	工程学（6，2.4%）

从三类装置在各学科上的份额分布来观察学科交叉情况，同步辐射光源装置具有相对平衡的学科交叉结构，化学、物理学、交叉科学、生物与生化、材料科学占比均高于 10%；相反，大装置 Y 的学科交叉结构平衡性最弱，仅物理学一类的占比就高达 68%，其余学科类分布则较为零散。这一现象的产生可能有两种原因：一是大装置 Y 装置目前主要作为研究物质特性、发现新现象的手段，应用于光物理、凝聚态物理、原子核物理等领域中，而同步辐射光源和散裂中子源的应用则相对更广；二是大装置 Y 装置是本世纪近来激光技术及超快科学的新突破，当前世界范围内成熟的装置屈指可数，设施的不完善导致了学科结构的偏斜。

在子学科层面（WoS 学科分类）结果如表 3 所示，针对大科学装置的相关子学科开展细粒度的分析，能够更好地展现装置与装置之间的学科关联程度，在已获取的装置论文数据中，散裂中子源装置共涉及 108 种小学科，同步辐射光源装置涉及 131 种小学科，大装置 Y 装置涉及 38 种小学科。各装置子学科分布 TOP10 及其占比、数量方面，材料科学-交叉学科、凝聚态物理学、化学物理学是散裂中子源装置最主要的相关子学科，生物化学与分子生物学、材料科学-交叉学科、化学-交叉科学是同步辐射光源最主要的相关子学科，光物理学、原子、分子和化学物理是大装置 Y 最主要的相关子学科。三类大科学装置的相关子学科存在较密集的重叠。

表 3 三类装置子学科分布 Top10

	散裂中子源			同步辐射光源			大装置 Y		
	学科名	论文数量	占比	学科名	论文数量	占比	学科名	论文数量	占比
1	材料科学,交叉学科	4752	17.23%	生物化学与分子生物学	2141	11.32%	光学	99	22.05%
2	凝聚态物理	4273	15.50%	材料科学,交叉学科	2007	10.61%	流体及等离子体物理学	59	13.14%
3	化学物理学	3824	13.87%	化学，交叉学科	1728	9.14%	应用物理学	44	9.80%
4	应用物理学	2692	9.76%	化学物理学	1398	7.39%	化学物理学	32	7.13%
5	化学，交叉学科	1720	6.24%	交叉科学	1319	6.97%	流体与等离子体物理学	25	5.57%
6	物理学，交叉学科	1149	4.17%	应用物理学	1223	6.47%	材料科学,交叉学科	24	5.35%
7	流体及等离子体物理学	1145	4.15%	凝聚态物理	842	4.45%	物理学，交叉学科	19	4.23%
8	化学,无机及核	701	2.54%	生物物理学	830	4.39%	仪器与仪表	17	3.79%
9	晶体学	575	2.09%	晶体学	730	3.86%	核科学与技术	17	3.79%
10	仪器与仪表	566	2.05%	细胞生物学	599	3.17%	纳米科学与纳米技术	15	3.34%

为直观展现三类科学装置的学科知识分布，基于各类装置子学科的分布情况构建装置-子学科分布关系（图 8）。在该图中，颜色越深的连线代表某一子学科在该装置论文中的占比越大，右侧颜色越深的节点代表该节点承载了越多三种装置给予的学科权重。

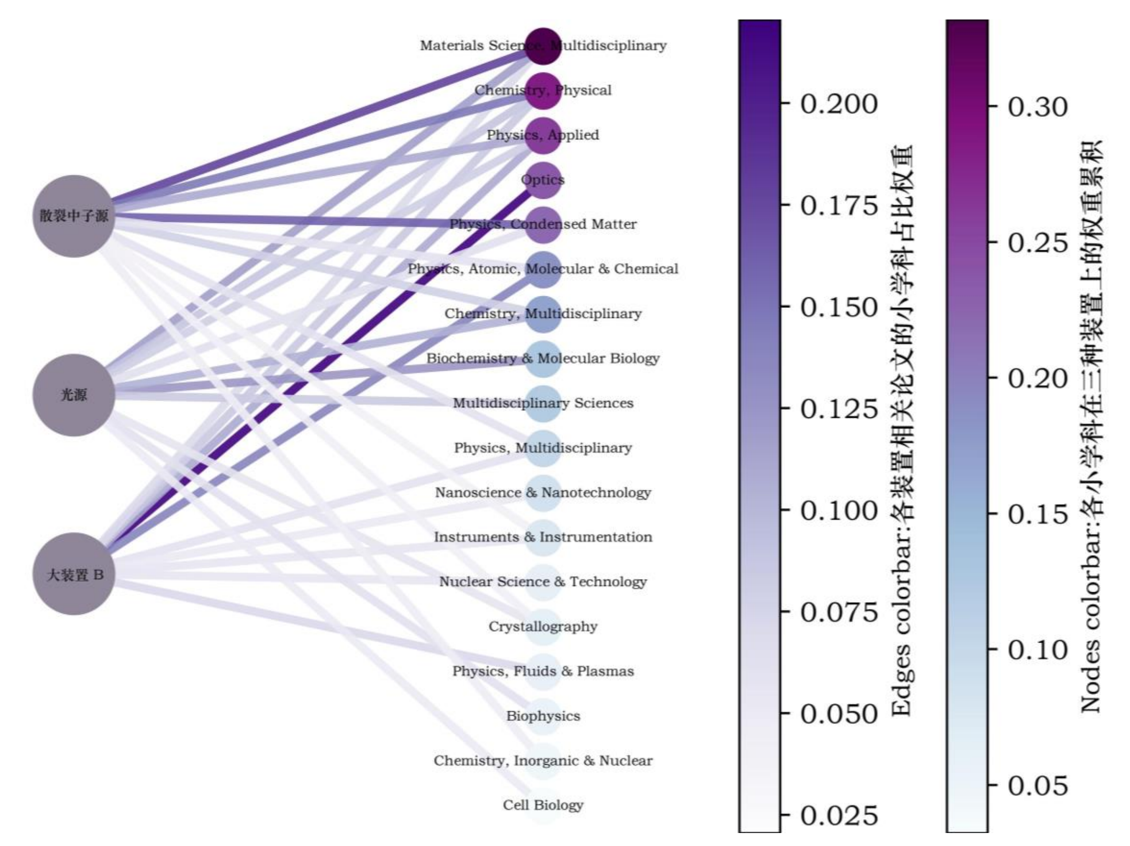


图 8　装置-子学科 分布关系

由图 8 可知，材料科学-交叉学科、化学物理学、应用物理学是三类科学装置重叠最密集的子学科；除此之外，凝聚态物理学、化学-交叉科学、晶体学等也是散裂中子源和同步辐射光源科学重叠较为密集的子学科，而原子、分子和化学物理、物理-交叉科学、仪器与仪表是散裂中子源和大装置 Y 科学装置存在较多重叠的子学科，以上学科在支撑科学装置集群相关学科发展时需要给予特别关注，同时，研究范围涉及以上学科的相关优势研究力量可以考虑作为支撑研究力量。

4.3 装置关联情况分析

为更量化地探讨三类装置间的相似性与差异性，下一步分别基于科学装置的学科分布向量和科学装置的论文网络两个角度展开分析。

在学科分布向量方面，表 4 为三类科学装置学科分布情况的两两余弦相似度比较结果，散裂中子源与同步辐射光源科学装置的学科分布相似度达 0.738，同时联系表 2 可知同步辐射光源科学装置具备有散裂中子源并不主要涉及的生物化学与分子生物学（占比高达 11.32%），和生物物理学（占比 4.39%），细胞生物学（3.17%）等，这说明散裂中子源科学装置与光源科学装置既有适度的学科分布相似度，又有各自的主要学科领域，从科研产出的角度来看两者共同布局是较为合理的。而大装置 Y 装置与其他两者的学科分布相似度中低，分别为 0.469 和 0.406，但其具有另两者所不具备的主要相关学科——光学（22.05%）、流体及等离子体物理学（5.57%）、核子学及技术（3.79%）、纳米科学与技术（3.34%），这说明从科研论文角度来看，大装置 Y 与散裂中子源、光源装置共同布局，可能起到一定的学科补充。

表 4　三类装置学科分布相似度

	散裂中子源	同步辐射光源	大装置 Y
散裂中子源	1.000	0.738	0.469
同步辐射光源	0.738	1.000	0.406
大装置 Y	0.469	0.406	1.000

透过直引网络和引文耦合网络中的簇分布，可以直观地展现不同装置论文间研究主题的关联程度，从而间接地展现装置间的联系。图 9 和图 10 是对三类科学装置产出论文的引文网络的分析结果，图 9 揭示了三类装置因论文间相互引用而产生的网络关系，图 10 揭示了三类装置因论文的参考文献耦合而产生的网络关系。两图呈现出类似的结论：一是散裂中子源（橙色节点）与同步辐射光源（紫色节点）装置相关论文之间存在着联系密集的簇团，而同步辐射光源另存有一块独立簇团，这说明两者的相关产出论文在一些领域中具有很强的关联性与同步性；同时，同步辐射光源装置具有散裂中子源相关论文较少涉及的较大型领域，即生物化学与分子生物学等。二是大装置 Y（绿色节点）装置的相关论文在网络图中处于边缘位置，这与大装置 Y 目前在世界范围内成熟装置少、出版物数量少有关，就该类装置目前已有论文来看，其与另两类装置的产出论文间的联系较为稀疏。

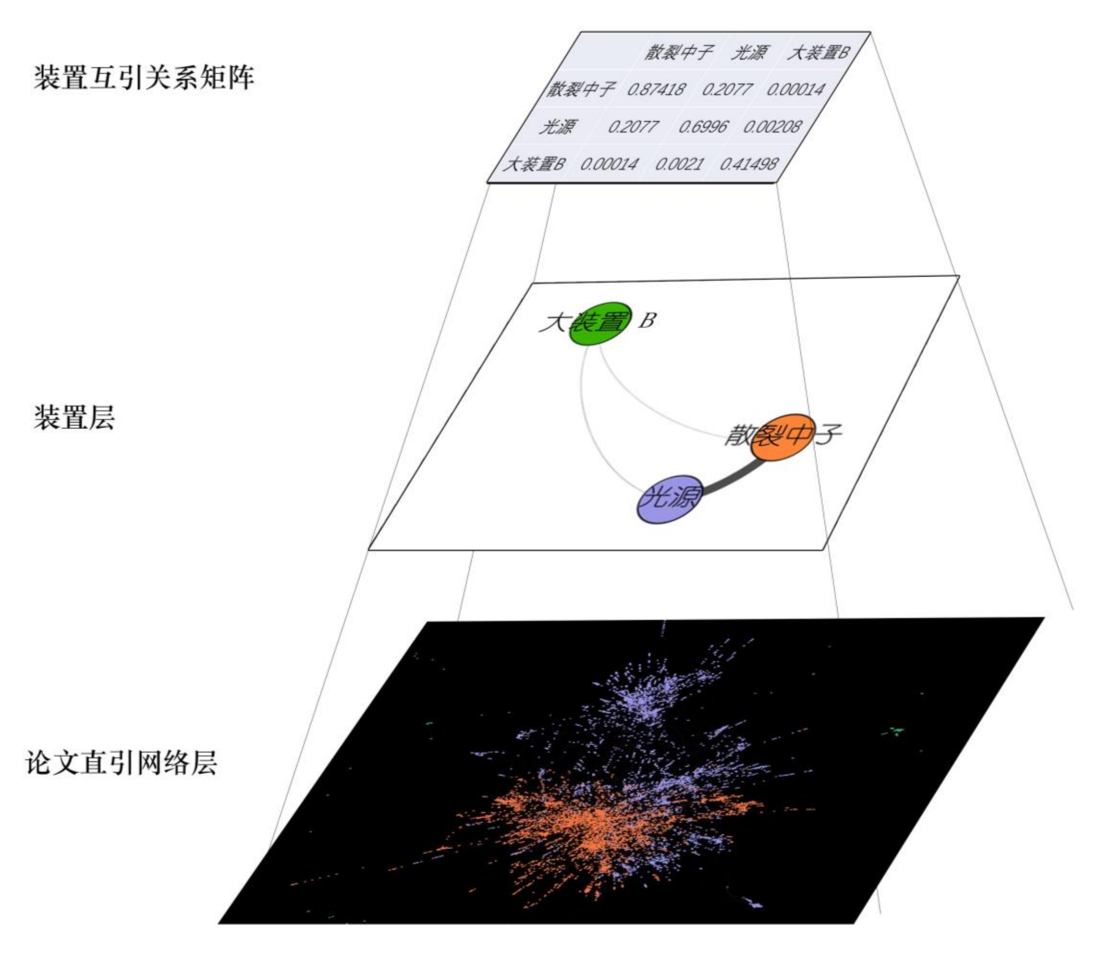


图 9 装置-论文直引分层网络

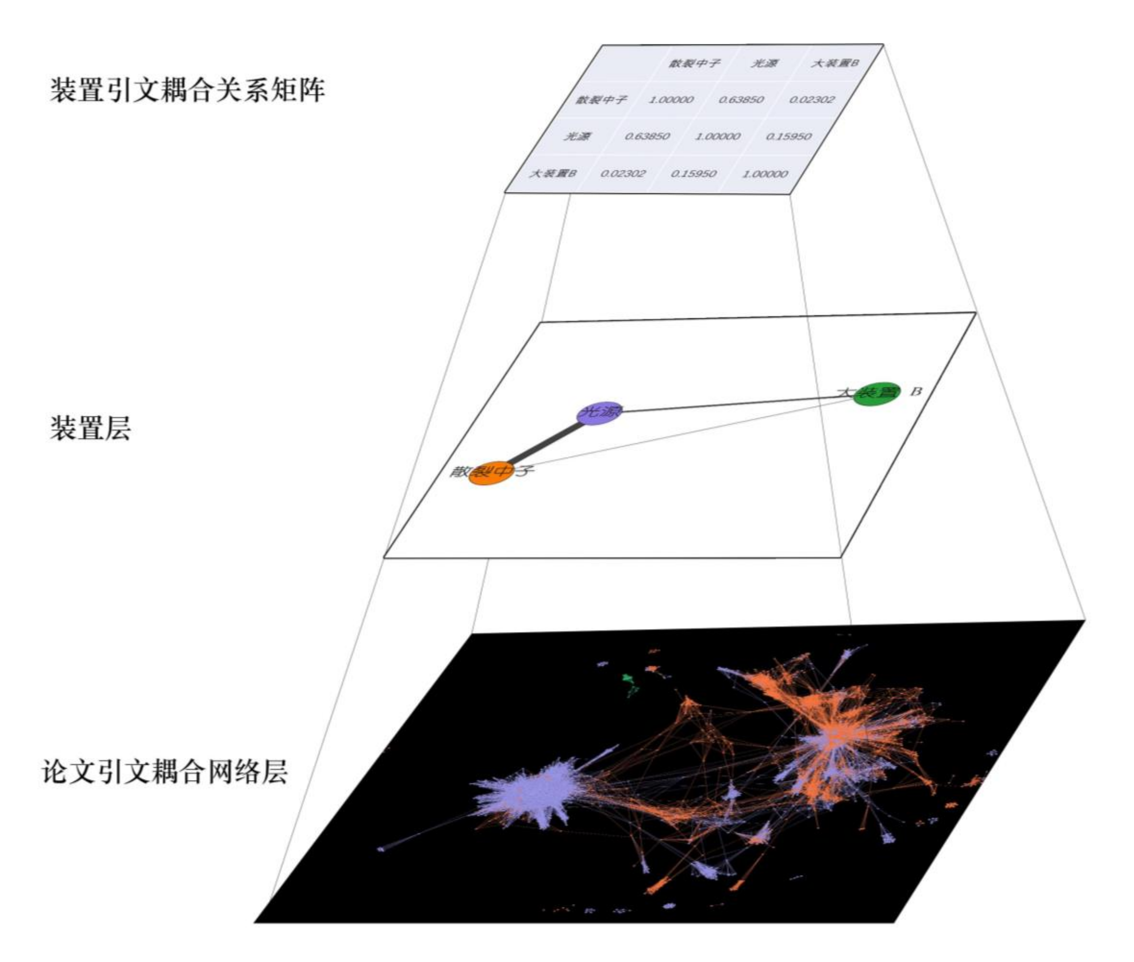


图 10 装置-论文引文耦合分层网络

上述数据结论从量化角度得出以下观点：散裂中子源、同步辐射光源、大装置 Y 三类装置均对基础性学科具有重大支撑作用，并均具有不同程度的交叉学科发展潜力；三类科学装置存在相同或相似的相关子学科，包括化学物理学、应用物理学、凝聚态物理学、晶体学、原子、分子和化学物理、仪器与仪表以及部分交叉科学，未来共同布局时需要给予特别关注；散裂中子源和同步辐射光源具有较强的学科分布和研究主题相似性，同步辐射光源装置对散裂中子源具有较为明显的学科和研究主题补充效应；大装置 Y 当前与另两者具有相对较低的学科关联度和研究主题相似性，但其对光物理学领域、原子分子和化学物理领域具有不可忽视的作用。

总体而言，数据层面上，散裂中子源与同步辐射光源装置关联度高，存在可能的互补关系，大装置 Y 与另两者关联度则较低，如要布局建议还需要进一步专家论证。

5 研究不足与展望

本研究从定量角度提出一种用于辅助大装置集群规划布局的通用型分析框架，以广东松山湖科学城为例，收集已有大装置与规划布局的大装置在国际上同类型装置的论文数据，分析其学科层面的内在关联，用以辅助判断规划布局的装置与已有装置的匹配关联程度。

本研究也存在一定局限：1）该方法暂仅从定量层面解读数据、探究大装置之间的学科关联程度，暂未涉及定性分析，可能会在准确性上存在一定误差；2）

本研究暂时仅从学科关联层面分析大装置之间的匹配程度，实际上大装置之间的匹配会涉及更多的要素（研究层面、地理层面、经济产业层面、人才层面等），学科关系仅是研究层面其中重要的一种；3）本研究仅使用了依托大装置产出的论文作为分析数据，有待拓展更多的分析资源使结果更全面；4）本研究仅使用已有的数据作分析，在结果呈现上会是对过去的总结，而大装置还具有一定的发展潜力，在装置集群不断扩充后有可能会产生新的用途；5）本研究在对即将规划布局的大装置及布局时间较短产出较少的大装置的分析上，使用的是世界同类型大装置的论文产出情况来替代分析目标，同类型的装置可能也会由于能量大小细微差异、集群装置配置等因素与分析对象的具体情况与实际分析对象存在一定的偏差，同时也由于不同装置的数据可获得性差异，在光源装置分析时舍弃了相对更适合作为替代分析的装置，可能也会对准确性存在影响。

下一步我们将进一步完善方法，挖掘大装置集群关系中更多可定量化分析的要素，拓展分析数据，结合定性定量分析，使分析结果对于大装置的规划布局更具有参考价值，从而为大装置群的功能协调与未来发展提供预见性、全局性的优化方案。

值得说明的是，科学计量学方法与其他方法体系一样，具有局限而无法包罗万象。本研究主要从装置产出论文的角度来量化窥探装置间的联系，只可为决策者提供一种客观的数据视角和辅助参考，在大科学装置集群的规划方面，更重要的是国家整体的规划和区域对大科学装置的需求，本方法需要融合大科学装置的内在原理、运作流程以及专家知识，才可在规划布局中发挥真正的作用。

参考文献

[1] 王济昌. 现代科学技术名词选编[M]. 郑州: 河南科学技术出版社, 2006.

[2] 梁永福, 盘思桃, 林雄等. 大科学装置集群的协同创新与产业带动效应——以广东大科学中心为例[J]. 科技管理研究, 2018, 38(03): 5-10.

[3] 肖国青, 李振中. 兰州重离子研究装置[J]. 中国科学院院刊, 2009, 24(01): 97-101+2+105.

[4] 齐芳: 国之重器奠定创新未来——十八大以来我国大科学装置成就综述[N], 光明日报, 2017.

[5] 陈套, 冯锋. 大科学装置集群效应及管理启示[J]. 西北工业大学学报(社会科学版), 2015, 35(01): 61-66.

[6] 中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要[EB/OL]. (2021-08-24) [2021-08-24]. http://www.gov.cn/xinwen/2021-03/13/content_5592681.htm.

[7] 黄振羽. 大科学装置与粤港澳大湾区发展的关系分析[J]. 科技管理研究, 2019, 39(18): 69-72.

[8] 张玲玲, 赵明辉, 赵道真等. 依托大科学装置的科学园空间布局及对策研究——以散裂中子源为例[J]. 工程研究-跨学科视野中的工程, 2019, 11(04): 338-348.

[9] 陈岸明, 魏东原. 粤港澳大湾区重大科技基础设施布局的优化分析——基于国际比较的视角[J]. 国际经贸探索, 2020, 36(10): 86-99.

[10] 程晓舫, 唐磊, 夏依林. 大科学装置共建共享及其对区域一体化的影响——以长三角为例[J]. 科技管理研究, 2020, 40(22): 26-31.

[11] 张玲玲, 赵明辉, 曾钢等. 文献计量视角下依托大科学装置的学科主题与合作网络研究——以上海光源为例[J]. 管理评论, 2019, 31(11): 279-288.

[12] Florio M, Sirtori E. Social benefits and costs of large scale research infrastructures[J]. Technological Forecasting and Social Change, 2016, 112: 65-78.

[13] Florio M, Forte S, Sirtori E. Forecasting the socio-economic impact of the Large Hadron Collider: A cost-benefit analysis to 2025 and beyond[J]. Technological Forecasting and Social Change, 2016, 112: 38-53.
[14] Michalowski S. The Impacts of Large Research Infrastructures on Economic Innovation and on Society: Case Studies at CERN[R]. Europe: Organisation for Economic Co-operation and Development (OECD), 2014.
[15] Hallonsten O. How expensive is Big Science? Consequences of using simple publication counts in performance assessment of large scientific facilities[J]. Scientometrics, 2014, 100(2): 483-496.
[16] Hallonsten O. Use and productivity of contemporary, multidisciplinary Big Science[J]. Research Evaluation, 2016, 25(4): 486-495.
[17] Heidler R, Hallonsten O. Qualifying the performance evaluation of Big Science beyond productivity, impact and costs[J]. Scientometrics, 2015, 104(1): 295-312.
[18] Qiao L L, Mu R P, Chen K H. Scientific effects of large research infrastructures in China[J]. Technological Forecasting and Social Change, 2016, 112: 102-112.

作者贡献声明：

冯凌子：提出研究思路，设计研究方案，采集与分析数据，撰写与修订论文；
张汝昊：设计研究方案，采集、清洗和分析数据，撰写与修订论文；
冯凯悦：提出研究命题，分析案例选择，修订论文；
袁军鹏：提出研究思路，论文修订。